

PROBLEMS AND RECENT ADVANCEMENTS IN EFFICIENT IMAGE SUPER-RESOLUTION

In recent years, as the amount of visual content (especially images) on the internet has grown, the need for enhancement methods has increased. Such methods, often called 'restoration methods', include image super-resolution, denoising, deblurring, artifact removal, etc. However, while the accuracy of such methods shows promising results, their efficiency still needs improvements. Often, the proposed methods are hardly applicable in edge devices and real-time applications.

Image Super-Resolution (ISR) is the dominant area among all restoration methods. The objective is to enhance the resolution of images from low-resolution (LR) to high-resolution (HR) format, aiming to recover finer details and visual quality. It finds its application in medical imaging, satellite observation, security surveillance, multimedia, etc. The solutions cover a wide range of computer vision methods, from classical approaches to deep learning methods, including Convolutional Neural Networks (CNNs), Generative Adversarial Networks (GANs), and Transformers.

Classical methods, such as bicubic interpolation, often fail to deliver visually pleasing results as they do not account for complex high-frequency details, leading to blurry and unrealistic images. On the other hand, CNNs are able to capture complex patterns through hierarchical feature extraction. However, higher accuracy is often achieved by stacking multiple convolutional layers, increasing the number of parameters, and Floating-point operations per second (FLOPS). As a result, deploying these large CNN models for SR on resource-limited devices becomes inefficient, highlighting the need for more efficient approaches. To solve this problem, for example, the authors of BSRN [1] proposed to use Blueprint Separable Convolution (BSConv) to reduce redundant operations by leveraging intra-kernel correlations.

Recently, Transformers have gained popularity in Computer Vision tasks and Image Restoration in particular. The architecture allows the model to capture long-range dependencies and better understand contextual information across image regions, showing better performance than CNNs. However, Transformers are known for their high computational cost due to self-attention mechanisms. With the increase of the input size, the computational cost of the attention increases quadratically, which is inefficient in ISR task. To solve this problem, localizing attention methods can be used. For example, SwinIR [2] limits self-attention to

local windows rather than the entire image, reducing computational costs. ESRT [3] uses a similar approach, but reduces computational cost with splitting factor of the inputs of self-attention (queries, keys and values).

The results of the challenge on Efficient Super-Resolution [4] show some other techniques to reduce computational cost and the size of the model, e.g. network pruning, re-parameterization, parameter-free attention mechanism, etc.

Analyzing the keywords and the semantics of the titles and abstracts of almost 3900 papers in ISR task in arXiv database (Fig. 1), it can be seen that the trend on efficient method for ISR is continuously growing, especially in transformer-based methods.

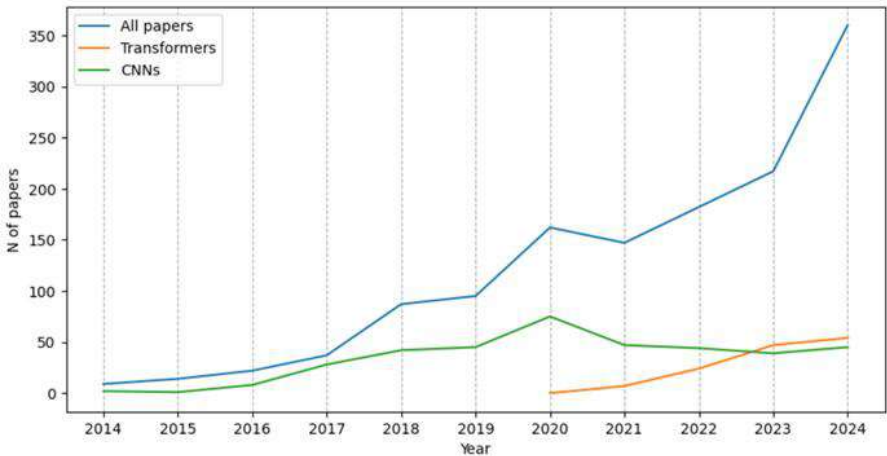


Figure 1 - The number of papers in ISR tasks that aim to improve efficiency/reduce memory usage or computational cost.

In conclusion, the field of efficient ISR is rapidly advancing, thanks to innovations in deep learning methods. These advancements are mostly driven by the demand for real-time performance on devices with limited resources and the development of lightweight architectures can be achieved through the optimization of the architecture and basic operations. However, challenges remain in computational efficiency and real-time application, making ISR a promising area for continued research and development.

REFERENCES

1. Z. Li, Y. Liu, X. Chen, H. Cai, et. al., Blueprint Separable Residual Network for Efficient Image Super-Resolution, Proceedings of the IEEE/CVF

Conference on Computer Vision and Pattern Recognition (CVPR) Workshops, 2022, pp. 833-843;

2. J. Liang, J. Cao, G. Sun, K. Zhang, L. Van Gool, R. Timofte, SwinIR: Image Restoration Using Swin Transformer, IEEE/CVF International Conference on Computer Vision Workshops (ICCVW), 2021, pp. 1833-1844;

3. Z. Lu, J. Li, H. Liu et. al., Transformer for Single Image Super-Resolution, IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2022, pp. 456-465;

4. B. Ren, Y. Li, N. Mehta, et. al., The ninth NTIRE 2024 efficient super-resolution challenge report, Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, 2024, pp. 6595-6631.